Estimation of Instantaneous Frequency Fluctuation in a Fast DVFS Environment Using an Empirical BTI Stress-Relaxation Model

Chen Zhou Xiaofei Wang Weichao Xu *Yuhao Zhu *Vijay Janapa Reddi Chris H. Kim

Department of Electrical & Computer Engineering
University of Minnesota
200 Union Street SE, Minneapolis, MN 55455, USA
zhoux825@umn.edu
*Department of Electrical & Computer Engineering
University of Texas at Austin
2501 Speedway, Austin, TX 78712, USA

***Abstract*—This work proposes an empirical Bias Temperature Instability (BTI) stress-relaxation model based on the superposition property. The model was used to study the instantaneous frequency fluctuation in a fast Dynamic Voltage and Frequency Scaling (DVFS) environment. V_{DD} and operating frequency information for this study were collected from an ARM Cortex A15 processor based development board running an Android operating system. Simulation results show that the frequency peaks and dips are functions of mainly two parameters: (1) the amount of stress or recovery experienced by the circuit prior to the V_{DD} switching and (2) the frequency sensitivity to device aging after the V_{DD} switching.**

Keywords – bias temperature instability; superposition property; dynamic voltage and frequency scaling; frequency degradation

I. Introduction

Modern processors are capable of throttling the chip power and performance whenever computing demand is low through techniques such as dynamic voltage and frequency scaling (DVFS) [1], Turbo/NVT (Near-threshold voltage) mode operation, and power gating. However, the recoverable BTI component during fast voltage transients can result in vastly different reliability behavior as compared to fixed voltage systems. Fig. 1 illustrates the frequency fluctuation in a DVFS environment based on real V_{DD} and frequency traces collected from an ARM Cortex A15 processor (see Fig. 11 for test setup). V_{DD} is chosen from several values based on computing demand, while clock frequency is decided by a V_{DD}-frequency lookup table. To calculate the threshold voltage (V_T) shift under continuous varying V_{DD}, this work proposes a BTI stress-relaxation model derived from reaction-diffusion (R-D) model [2] and superposition property assumption.

Since BTI induced V_T shift slows down logic circuit speed, actual maximum logic path frequency should be slower than the fresh frequency that is before aging. When the supply switches from a low V_{DD} to a high V_{DD}, the frequency instantaneously spikes due to the large recovery that took place in the preceding low V_{DD} mode. Since the clock frequency is set by the worst case long term aging, a sufficient guard-band exists right after the switching occurs. In contrast, when the processor switches back to a low V_{DD} mode, the large degradation in the preceding high V_{DD} mode results in a large dip in the circuit frequency. This implies the clock frequency has to be low enough to account for the worst case dip in logic path frequency which occurs after a long DC stress period in high V_{DD} mode followed by a low V_{DD} operation. Although there have been previous studies exploring the influence of workload on BTI [4-5], no prior work has analyzed the frequency fluctuation effect described above in realistic DVFS environment.

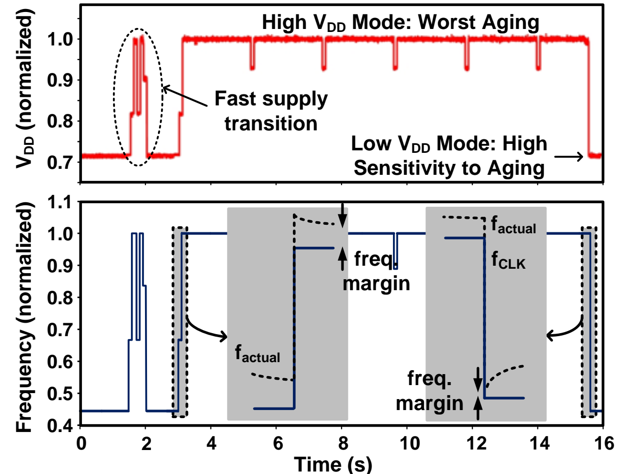


Fig. 1. V_{DD} and frequency traces for an ARM Cortex-A15 processor. Fast BTI coupled with the sensitivity difference between high V_{DD} and low V_{DD} modes generate freq. peaks and dips during DVFS operation.

II. BTI Stress-Relaxation Model

To study the impact of fast voltage switching on the actual frequency shift, we firstly develop an empirical BTI stress and relaxation model based on the superposition property described in Fig. 2. Degradation result from a complex DVFS voltage trace can be calculated by

accumulating the individual aging effects of short stress or relaxation voltage segments.

A. DC Stress-relaxation Model

For single V_{DD} stress mode, we use the traditional R-D t^n model where $\Delta V_T(t)$ is expressed as:

$$\Delta V_T(t) = A \cdot exp(\alpha V_{DD} - \frac{E_a}{kT}) \cdot t^n \quad (1)$$

Here, n is the stress time exponent, α is the voltage acceleration factor, E_a is the thermal activation energy, k is the Boltzmann constant, and T is the absolute temperature.

Based on superposition assumption, the stress-relaxation curve due to a long voltage pulse can be represented as the sum of those due to two short individual voltage pulses, as shown in Fig. 2. Therefore, the relaxation formula is determined by this requirement:

$$\Delta V_T(t) = A \cdot exp\left(\alpha V_{DD} - \frac{E_a}{kT}\right) \cdot [t^n - (t - t_{s0})^n] \quad (2)$$

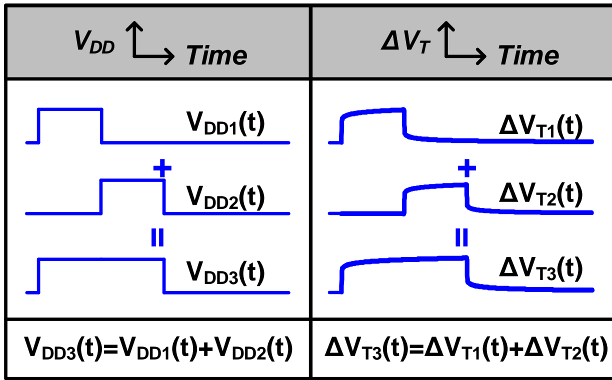


Fig. 2. Our empirical model assumes that BTI induced V_T shift for a voltage sequence can be expressed as the superposition of V_T shifts for the individual stress-relax voltage segments.

Here, t_{s0} is the pulse duration given in Fig. 3. Parameters used in the model such as α, E_a, n are extracted from experimental data of a previous test chip capable of separately monitoring PBTI and NBTI shifts (Fig. 4). Since a complicated V_{DD} trace can always be segmented into series of short V_{DD} pulses, the overall V_T shift curve is the sum of individual stress-relaxation curve.

B. AC Stress-relaxation Model for 50% Duty Cycle

The proposed model can be further extended for AC stress by accounting for the cycle-by-cycle stress and relaxation effects. Fig. 5 demonstrates the cycle-by-cycle behavior of BTI induced V_T shift for NMOS and PMOS in an inverter driven by a 1GHz input with a 50% duty-cycle. Each transistor is in relaxation phase right before being switched on, indicating that the bottom boundary of each zigzag curve represents the effective V_T shift for long term AC stress prediction. By applying the proposed model, the V_T shift before each transition point can be calculated as derived in Fig. 6.

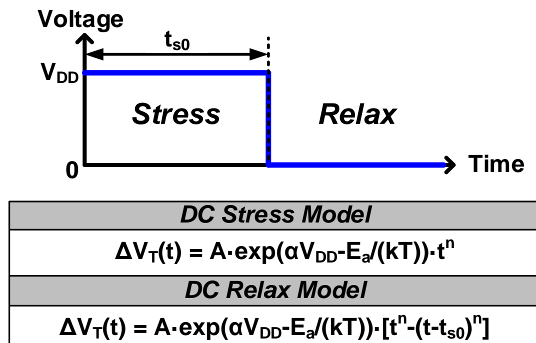
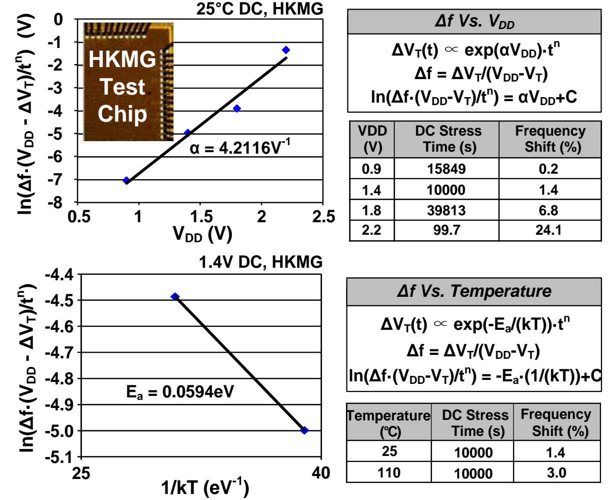


Fig. 3. DC stress and relax models.



VDD (V)	DC Stress Time (s)	Frequency Shift (%)
0.9	15849	0.2
1.4	10000	1.4
1.8	39813	6.8
2.2	99.7	24.1

Temperature (°C)	DC Stress Time (s)	Frequency Shift (%)
25	10000	1.4
110	10000	3.0

Fig. 4. Voltage and temperature acceleration factors for the proposed BTI models are extracted from experimental data.

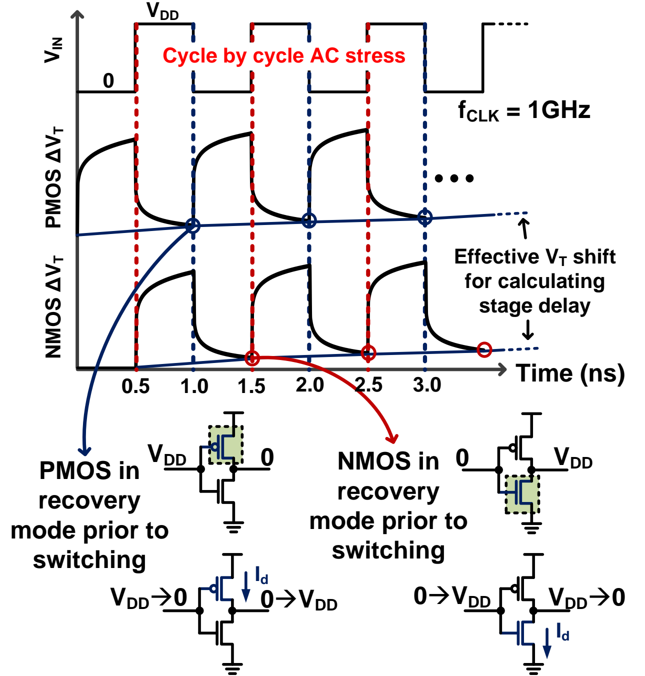


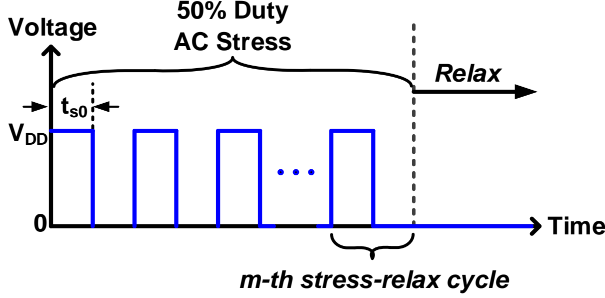
Fig. 5. Illustration of cycle-by-cycle AC stress. In digital circuits, a transistor is always in a recovery mode prior to turning on and hence the post recovery V_T shift must be used to calculate the actual delay degradation.

After m cycles of 50% duty AC stress with pulse duration time t_{s0}, the ΔV_T is expressed as:

$$\Delta V_T(m) = A \cdot exp\left(\alpha V_{DD} - \frac{E_a}{kT}\right) \cdot \sum_{i=1}^{2m} (-1)^i (i \cdot t_{s0})^n \tag{3}$$

ΔV_T recovery after end of whole 50% duty AC stress process which has m cycles in total is modeled as:

$$\Delta V_T(t) = A \cdot exp\left(\alpha V_{DD} - \frac{E_a}{kT}\right) \cdot \sum_{i=0}^{2m-1} (-1)^i (t - i \cdot t_{s0})^n \tag{4}$$



AC Stress Model
ΔV_T (at the end of m-th stress-relax cycle) =
$A \cdot exp(\alpha V_{DD} - E_a/(kT)) \cdot \sum_{i=1,2,...2m-1,2m} (-1)^i (i \cdot t_{s0})^n$
AC Relax Model
ΔV_T (recovery after m stress-relax cycles) =
$A \cdot exp(\alpha V_{DD} - E_a/(kT)) \cdot \sum_{i=0,1,...2m-2,2m-1} (-1)^i (t - i \cdot t_{s0})^n$

Fig. 6. AC stress and relax models. The overall aging is estimated by aggregating the aging effects of individual stress and relax period segments. The computation time for m=7.884×10^{15} (5 years, 1 GHz, 5% activity factor) was 0.054 sec using the Mathematica software.

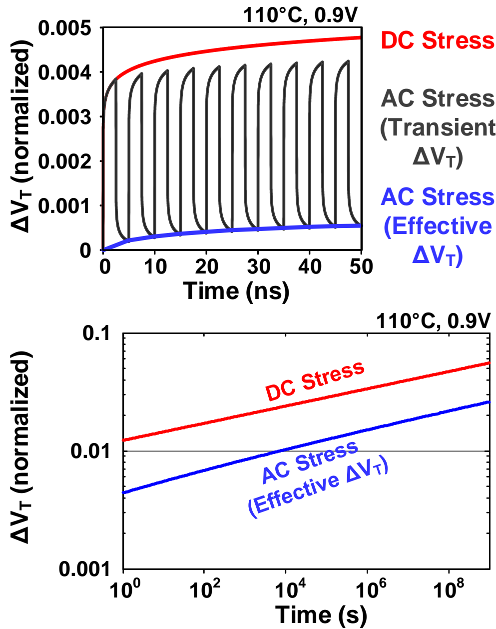


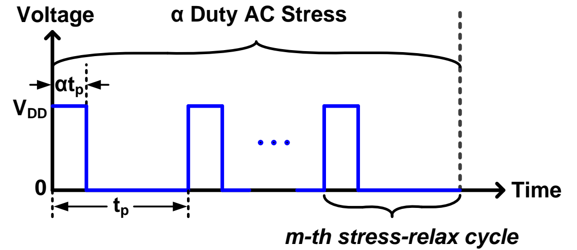
Fig. 7. ΔV_T under DC stress and 200MHz AC stress at 110°C. First 50ns in linear-linear scale (above) and 10 years in log-log scale (below).

The total amount of aging is basically the sum of all individual aging components for m stress-relaxation cycles. Since the m value can be extremely large even for a few seconds of run-time and a GHz clock frequency, we have to rely on a numerical method to calculate the overall sum. By using the integrated numerical functions in the Mathematica software, we were able to complete the calculation in 0.054s for a five-year lifetime. Using this method, V_T degradation for DC stress and 200MHz AC stress are calculated under a 0.9V, 110°C stress condition (Fig. 7). Result shows that AC stress induced V_T shift is about 40% of the DC case, while the time slope in log scale is approximately the same. These results are in line with the previous experimental observations [6], [7].

C. AC Stress-relaxation Model for an Arbitrary Duty Cycle

During actual circuit operation, transistors may experience AC stress with arbitrary duty cycles. The previous model can be extended to accommodate non-50% duty cycles as shown in the following equation and Fig. 8 where the superposition principle is applied to each clock cycle with asymmetric stress and relaxation time:

$$\Delta V_T(m) = A \cdot exp(\alpha V_{DD} - \frac{E_a}{kT}) \cdot \sum_{i=1}^{m} [-(i-\alpha)^n + i^n] \cdot t_p^n \tag{5}$$



α Duty AC Stress Model
ΔV_T (at the end of m-th stress-relax cycle) =
$A \cdot exp(\alpha V_{DD} - E_a/(kT)) \cdot \sum_{i=1,2,...m-1,m} [-(i-\alpha)^n + i^n] \cdot t_p^n$

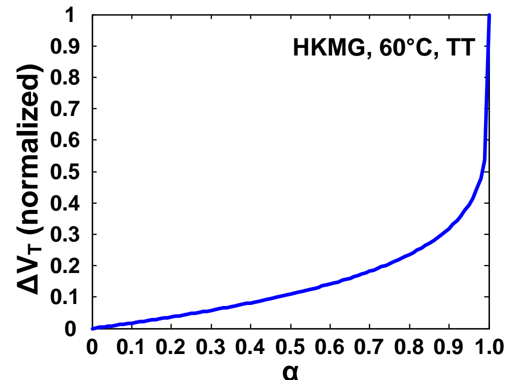


Fig. 8. AC stress and relax models with any duty cycle α, estimated by aggregating the aging effects of individual stress and relax period segments.(above) Duty α vs. ΔV_T (below).

Here, α denotes duty cycle, which ranges from 0 to 1. ΔV_T exhibits more or less a linear relationship for α values from 0 and 0.7, and then rapidly increases as α approachs 1

(i.e. DC stress). There is a discrepancy between the results in Fig. 8 and those reported in [5], which demonstrates a big change at $\alpha = 0$ but little change at $\alpha = 1$. The difference between the two results can be attributed to the different amount of recovery during relaxation phase. A large recovery, such as 80% will induce a sudden rise around $\alpha = 1$ while a weak recovery, such as 50% will result in an rise only at $\alpha = 0$ which is expected. The measurement data from our test chip shows a large recovery resulting in the trend shown in Fig. 8. Since the amount of recovery is affected by many factors such as the thickness of oxide layer and stress frequency [8], the exact dependency of ΔV_T on duty cycle α may vary from technology to technology.

III. Estimation of Frequency Fluctuation under DVFS

Frequency degradation of a given circuit path can be estimated using the proposed BTI model and look-up maps in Fig. 9 for translating V_T shifts to delay shifts.

A. Delay Sensitivity Map

The dependence of an inverter's pull-down and pull-up delays (normalized value) on the threshold voltage shift and supply voltage were simulated using HSPICE and displayed in Fig. 9. These delay maps can be used to translate the V_T shift calculated using the proposed BTI models, to delay degradation for a given circuit path. For example, the overall delay degradation of a ring oscillator circuit under DC stress can be calculated by summing up the pull-down and pull-up delays of each stage (considering those with fresh devices) using the two delay maps in Fig. 9. Any mismatch between the pull-up and pull-down delays resulting from P/N skew and time-zero V_T difference can be captured using two separate delay sensitivity maps. Notice that the sensitivity to aging increases as the V_{DD} is reduced.

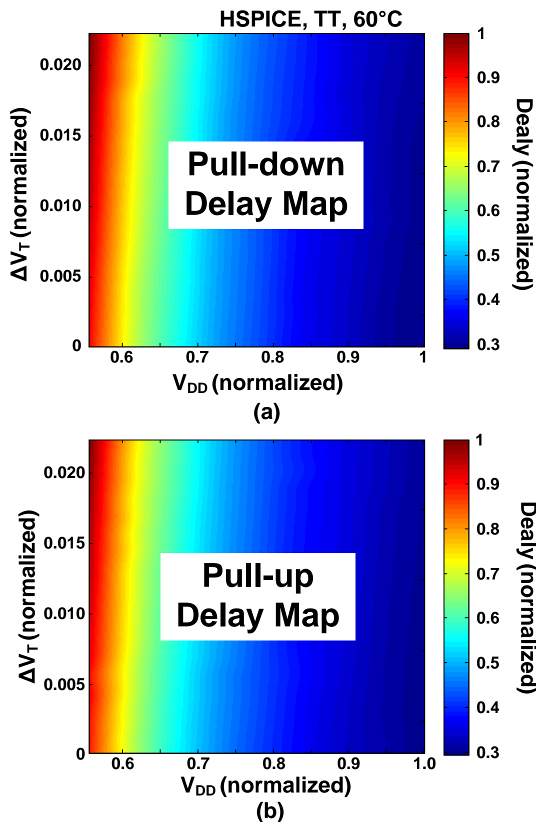


Fig. 9. (a) Pull-down delay map and (b) pull-up delay map for translating V_T shift to delay degradation at different V_{DD}'s.

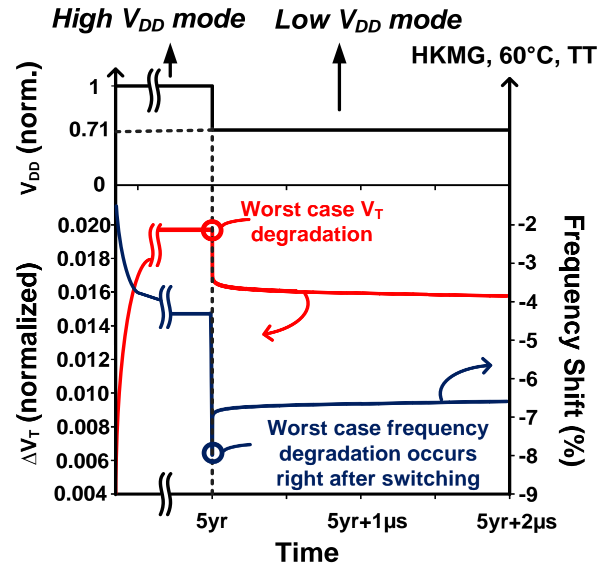


Fig. 10. Worst case scenario for ΔV_T and frequency shifts. 5 years of DC stress in high V_{DD} mode followed by a low V_{DD} mode.

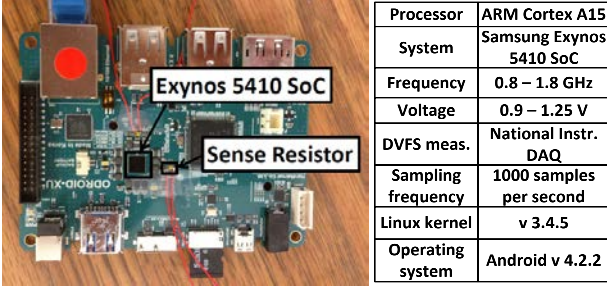
B. Worst-case Scenario Example

Fig. 10 shows the V_T shift and the corresponding frequency degradation for an extreme worst-case scenario, where the circuit is kept in a high V_{DD} DC stress mode for a long period of time (five years), and is suddenly switched to a low V_{DD} active mode. This scenario combines both the highest V_T shift and the highest aging sensitivity. The circuit under test for this simulation was a simple inverter chain path. However, the same methodology can be used to estimate the frequency shift of a more complex path composed of different types of gates (e.g. NAND, NOR, INV), different fan outs, and different duty cycles. The frequency degradation at the end of the long DC stress period in Fig. 10 is calculated as 4.5%, followed by a 7.64% sharp dip when switching to the low V_{DD} mode. The magnitude of the frequency drop from fresh frequency determines the maximum operating frequency of the processor in the low V_{DD} mode. After switching to a low V_{DD} mode, the frequency is recovered by 1% within 2μs due to BTI relaxation.

C. Frequency Fluctuation in Fast DVFS Environment

In order to study the actual frequency degradation in a real world DVFS system, the ARM Cortex A15 processor based Android system shown in Fig. 11 has been tested. The V_{DD} and operating frequency information can be collected in real time through the test setup. The V_{DD} level follows dynamic patterns while the system is used to process different tasks, e.g. navigating different websites and running benchmark applications, as illustrated in Fig. 12. The percentage frequency degradation can be readily predicted by

the proposed modeling method. As noted earlier, the worst



Processor	ARM Cortex A15
System	Samsung Exynos 5410 SoC
Frequency	0.8 – 1.8 GHz
Voltage	0.9 – 1.25 V
DVFS meas.	National Instr. DAQ
Sampling frequency	1000 samples per second
Linux kernel	v 3.4.5
Operating system	Android v 4.2.2

Fig. 11. Supply voltage and frequency trace measurement setup.

case frequency dip occurs when the supply voltage switches from high to low due to the combined effect of aggravated aging in high voltage mode and the higher sensitivity in low voltage mode. The magnitude of the frequency dip is a function of the time interval in the p receding high voltage mode and the voltage difference. Among the various benchmarks, the worst % frequency dip occurs for 3D-raytrace; Δf=1.0% at t =6s when V_{DD} drops by 29% after staying in high V_{DD} mode for 5.8s. In cases where the voltage drop is relatively small (e.g. t=6s for the website sina.com), a small initial dip in the percentage frequency shift (~0.02%) exists after which the recovery takes place to bring the frequency back towards the fresh state.

D. V_{DD} Ramp Time vs. Frequency Shift Dip

Based on computing demand, DVFS system changes V_{DD} and operating frequency. Previous simulations were run under the assumption that all changes happen immediately. In realistic case, both V_{DD} and operation frequency need transition time to complete the switch. These two conditions lead to different worst case degradation in terms of potential largest sensitivity.

Realistically, V_{DD} changes gradually from previous value to later value, thus can be modeled as a ramp voltage. The V_{DD} and frequency relation can be obtained from a preset V_{DD}-frequency look up table. Operation frequency must be kept below value from table to ensure no bit error. There are two approaches to adapt the frequency during the transition: (a) instant frequency change and (b) gradual frequency change. In approach (a), transition frequency is set to the lower frequency that prevents timing failure. Specifically, if the system is about to decrease V_{DD}, it switches operation frequency to lower target firstly, and then start to reduce V_{DD}. On the contrary, when system needs to switch back to higher V_{DD}, it switches frequency to higher target after V_{DD} switching is finished. The use of guard-band leads to the potential loss of computing power because circuits could have operated under a higher frequency in transition. To overcome this disadvantage, approach (b) tracks V_{DD} ramp voltage and maintains the highest frequency during transition. This method minimizes the potential loss in the computing power. For example, a VCO built in processor [9][10] keeps monitoring supply voltage and helps PLL generating maximum operation frequency without significant guard-band based on detected value.

Fig. 13 introduces a way to calculate frequency degradation at V_{DD} fall point and obtain more realistic result compared to those shown in Fig. 12. To simplify the analysis process, result shown in Fig. 12 is over pessimistically assumed that all V_{DD} and clock frequency switching happen instantaneously and no BTI recovery is absent during transition. As shown with red blocks in Fig. 13, transitional voltage is segmented into individual pulses and simulated based on the proposed model. Applying method shown in Fig. 13, Fig. 14 explores the likely worst scenario for normal operation. In this condition, supply voltage of ring oscillator is firstly kept at highest value for 5 years, a typical lifetime for general processor, and then switched to lowest V_{DD} through a ramp curve, while clock frequency tracks changing V_{DD} continuously. The relationship between V_{DD} ramp time, such as 0.01us, 0.1ms and 1ms, and BTI induced frequency degradation dip is demonstrated. Frequency degradation percentage dip always happens at the end of each transition because of the combined effect of recovering V_T shift and increasing aging sensitivity, both as the result of lower V_{DD}.

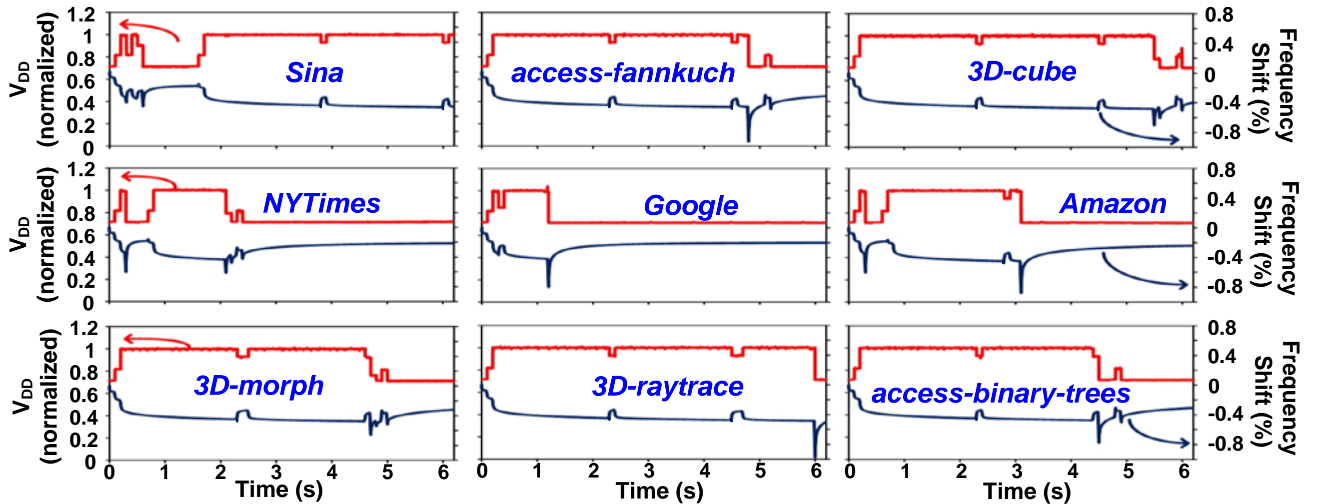


Fig. 12. Supply voltage (measured) and % frequency shift (simulated using our BTI models) of an ARM Cortex-A15 processor while navigating popular websites and running benchmark applications.

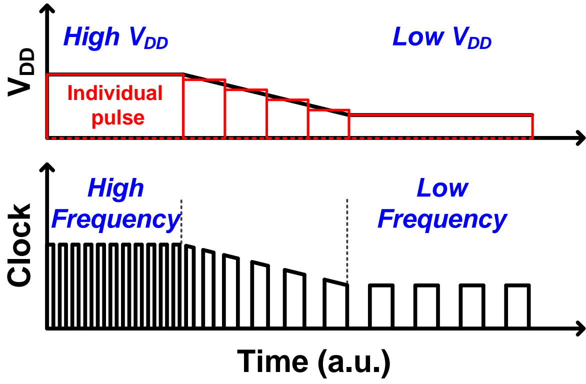


Fig. 13. Supply voltage and clock frequency switching from high V_{DD} mode to low V_{DD} mode.

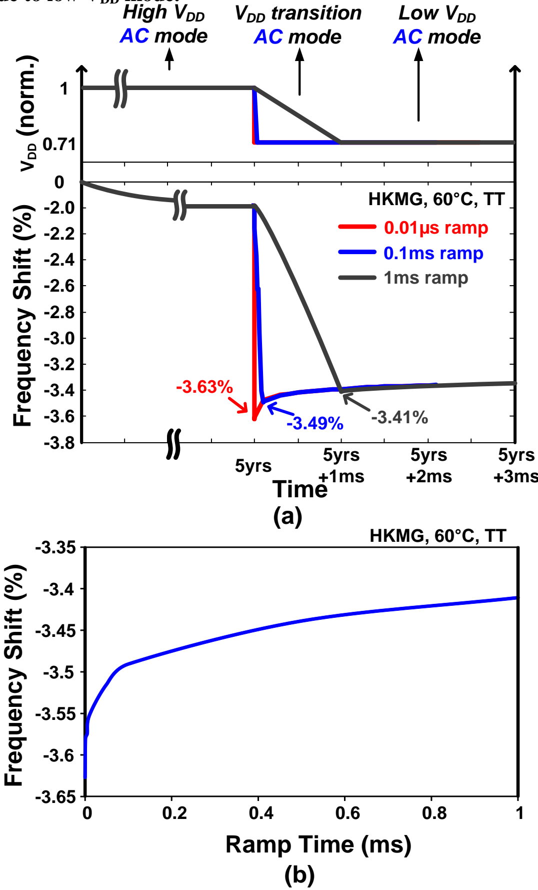


Fig. 14. (a) Frequency shift trace with V_{DD} switch ramp time of 0.01us, 0.1ms and 1ms. (b) Frequency shift dip vs. V_{DD} switch ramp time.

Lower ramp rate means longer time for V_T to recover before the highest aging sensitivity result is reached at lowest V_{DD}. Consequently, smaller V_T shift and less significant frequency degradation happens at the most sensitive point. As the ramp time increases from 0ms to 1ms, the frequency degradation dip reduces from -3.63% to -3.41%.

IV. Conclusions

Modern processor is designed with the ability to work under DVFS environment, which calls for the necessity to consider frequency shift fluctuation under continuous change of supply voltage compared to traditional single power supply and operating frequency case. In this work, we propose an empirical BTI model with superposition property that allows us to replace complicated BTI induced V_T shift calculation under changing V_{DD} with a summation of series simple calculation under single V_{DD}. This model shows the ability in dealing with real V_{DD} trace collected from ARM Cortex A15 processor running benchmark applications on an Android system. Simulation results show the frequency shift dips which happen when supply voltage switches from high V_{DD} mode to low V_{DD} mode limit operation frequency. Proposed model can also estimate BTI aging under various duty cycle and V_{DD} switch ramp time.

Acknowledgment

This work was supported by in part by the National Science Foundation (NSF) Award CCF-1255937 and the Semiconductor Research Corporation (SRC). The authors would also like thank Dr. Vijay Reddy at Texas Instruments for technical feedback and encouragements.

References

[1] J. Howard, S. Dighe, S.R. Vangal, G. Ruhl, N. Borkar, S. Jain, V. Erraguntla, M. Konow, M. Riepen, M. Gries, G. Droege, T. Lund-Larsen, S. Steibl, S. Borkar, V.K. De, R. Van Der Wijngaart, "A 48-Core IA-32 Processor in 45 nm CMOS Using On-Die Message-Passing and DVFS for Performance and Power Scaling," *IEEE Journal of Solid-State Circuits*, vol.46, no.1, pp. 173-183, Jan. 2011

[2] Alam, M.A., "A critical examination of the mechanics of dynamic NBTI for PMOSFETs," *IEEE International Electron Devices Meeting*, pp. 14.4.1,14.4.4, 8-10, Dec. 2003

[3] J.B. Velamala, K.B. Sutaria, H. Shimuzu, H. Awano, T. Sato, G. Wirth, Yu Cao, "Logarithmic modeling of BTI under dynamic circuit operation: Static, dynamic and long-term prediction," *IEEE International Reliability Physics Symposium*, pp. CM.3.1,CM.3.5, 14-18, April 2013

[4] Min Chen, H. Kufluoglu, J. Carulli, V. Reddy, "Aging sensors for workload centric guardbanding in dynamic voltage scaling applications," *IEEE International Reliability Physics Symposium*, pp.4A.2.1,4A.2.5, 14-18, April 2013

[5] E. Mintarno, V. Chandra, D. Pietromonaco, R. Aitken, R.W. Dutton, "Workload dependent NBTI and PBTI analysis for a sub-45nm commercial microprocessor," *IEEE International Reliability Physics Symposium*, pp.3A.1.1,3A.1.6, 14-18, April 2013

[6] A. Bansal, Kai Zhao, Jae-Joon Kim, R. Rao, "Bias Temperature Instability model for digital circuits-predicting instantaneous FET response," *IEEE International Reliability Physics Symposium*, pp.CR.2.1,CR.2.4, 10-14, April 2011

[7] D.P. Ioannou, K. Zhao, A. Bansal, B. Linder, R. Bolam, E. Cartier, J. Kim, R. Rao, G. La Rosa, G. Massey, M. Hauser, K. Das, J.H. Stathis, J. Aitken, D. Badami, S. Mittl, "A robust reliability methodology for accurately predicting Bias Temperature Instability induced circuit performance degradation in HKMG CMOS," *IEEE International Reliability Physics Symposium*, pp.CR.1.1,CR.1.4, 10-14, April 2011

[8] S.V. Kumar, C.H. Kim, S.S. Sapatnekar, "A Finite-Oxide Thickness-Based Analytical Model for Negative Bias Temperature Instability," *IEEE Transactions on Device and Materials Reliability*, vol.9, no.4, pp.537, 556, Dec. 2009

[9] A. Wang, S. Naffziger, "Adaptive Techniques for Dynamic Processor Optimization: Theory and Practice" pp. 124-142, 2008

[10] T. Fischer, J. Desai, B. Doyle, S. Naffziger, B. Patella, "A 90-nm variable frequency clock system for a power-managed Itanium architecture processor," *IEEE Journal of Solid-State Circuits*, vol.41, no.1, pp.218,228, Jan. 2006